

Module 11 Lab Activity: Core Concepts in Analysis of Variance
PSY 652 Research Methods
Nov 11, 2020

Description of the dataset:

The StressReduction.csv dataset includes the stress levels of 120 subjects who were randomly assigned to either no treatment or one of three stress reduction treatments (referred to as Programs). The variables are:

- Program = a text variable representing which of the four stress reduction treatments a subject was assigned to (None, Healthy Habits, Get Movin, and Recharge).
- ProgramCode = assigned a numeric value for program (0 = None, 1 = Healthy Habits, 2 = Get Movin, 3 = Recharge).
- StressLevel = reported stress level for each participant on a scale of 1-10, in which higher values represent higher stress levels.

1. Download the “StressReduction.csv” dataset from the Module 11 Lab module and save it into a project folder.
2. Create a new R notebook from your project file and name it “StressReduction_notebook”
3. Create a new R chunk with a first level header: “Load Libraries”
 - a. Load the psych & tidyverse packages
4. Create a new R chunk with a first level header: “Import Data”
 - a. Read in the “StressReduction.csv” dataset and assign it to an object named “stress”
5. Create a new R chunk with a first level header: “Factor the grouping variable”
 - a. Factor the ProgramCode variable so that it “None” takes on level 0, “Healthy Habits” takes on level 1, “Get Movin” takes on level 2 and “Recharge” takes on level 3. Name this new variable ProgramCode.f

Hint:

```
data <- mutate(data, Variable.f = factor(Variable, levels = c(0,1,2,3), labels = c("None", "Healthy Habits", "Get Movin", "Recharge")))
```

6. Create a new R chunk with a first level header: “Get descriptives”
 - a. Use any method to calculate descriptive statistics for the StressLevel variable (the mean you calculate in this step will represent the Grand Mean of StressLevel)
 - b. Use a combination of group_by() and summarize to get the mean value of StressLevel by ProgramCode.f
 - c. In the white space, write 1-2 sentences interpreting how the group means differ from other groups and from the grand mean.

7. Create a new R chunk with a first level header: “Create boxplots of StressLevel across treatment groups”
 - a. Use the ggplot function to create boxplots for StressLevel across the four treatment groups (ProgramCode.f). All four boxplots should be *in the same plot* (Not faceted).

Hint: `ggplot(dataframe, aes(y = outcome_variable, color = categorical_variable)) + geom_boxplot()`

- b. In the white space, write 1-2 sentences explaining what this plot shows.
8. Create a new R chunk with a first level header: “Conduct an ANOVA”
 - a. Use the lm function to build a model in which StressLevel is regressed on ProgramCode.f.
 - b. Use the anova function to display anova output for this model
 - c. Calculate the η^2 for the model (hint: $\eta^2 = SS_{\text{effect}} / SS_{\text{total}}$)
 - d. In the white space, answer the following questions:
 - i. Interpret the F statistic. Is there a significant difference between group means for program level?
 - ii. Interpret the model η^2
9. Create a new R chunk with a first level header: “Conduct planned contrasts”
 - a. In a new R chunk with a second level header: “Create contrasts.” In this code chunk, create *three contrasts* with pairwise comparisons between StressLevel in the no treatment versus three treatment groups. That is: None vs. Healthy Habits, None vs. Get Movin, and None vs. Recharge. Save each of these three contrasts to a new object.
 - b. In a new R chunk with a second level header: “Bind contrasts to predictor variable”
 - i. Use the contrasts and cbind functions to attach the three contrasts you just created to the predictor variable as a new attribute (i.e., ProgramCode.f). View this updated variable to make sure that your weights were assigned correctly as an attribute.

Hint:

```
# attach the contrasts to your variable:
contrasts(dataframe$variablename) <- cbind(contrast1, contrast2, contrast3)
# use attributes() to ensure the attribute was added correctly
attributes(dataframe$variablename)
```

10. Create a new R chunk with a second level header: “Run ANOVA with contrasts”

- a. Build an ANOVA model in which StressLevel is regressed on ProgramCode.f (to which the contrasts have been attached) and get model output using the `summary.lm` function

Hint: `summary.lm(modelname)`

11. In the white space, answer the following questions:

- a. What does the estimate for the intercept represent (hint, you calculated this before!)
- b. Interpret each of the three contrasts. Is the mean StressLevel for any of the treatment groups significantly different from the “no treatment” group?
- c. What is the difference between a planned contrast and a post-hoc analysis? When would you use one method over the other?